Contribution ID: **221**   Type: **Talk**

Using Machine Learning methods for improving data quality in the ALICE experiment

Thursday 2 August 2018 16:20 (20 minutes)

Data Quality plays an important role in many high-energy physics experiments, e.g. the ALICE experiment at the Large Hadron Collider (LHC), CERN. Currently used methods for quality assurance problems such as quality label assignment or particle identification, rely heavily on human expert judgments and complex computations. Those tasks, however, can be easily addressed by modern machine learning methods. In this talk, we present an overview of machine learning approaches to several tasks. The first task we address is automatic assignment of data quality label. Our results for the Time Projection Chamber (TPC) show that using the best performing algorithm, i.e. Random Forest, we can correctly classify over 75% of all data without any human interaction with over 95% precision. We also show how to use a Random Forest to improve the current approach for Particle identification task. Instead of manual 'cut-offs', we propose to select desired type of particles with more complex classification algorithms. Our tests indicate that with our solution we can distinguish up to 16.4% more of desired particles, while increasing the purity of resulting subsample by 9.33%. Finally, as a first step toward a semi-real-time anomaly detection tool, we present a proof-of-concept solution for generating the possible responses of detector clusters to particle collisions, using the real-life example of the TPC. Its essential component is a fast generative model that allows to simulate synthetic data points that bear high similarity to the real data, so they can be compared with the real detector output. Leveraging recent advancements in machine learning, we propose to use state-of-the-art generative models, namely Variational Autoencoders (VAE) and Generative Adversarial Networks (GAN), which are up to 103 faster than currently used GEANT3 simulation tool.

Author: GRACZYKOWSKI, Lukasz Kamil (Warsaw University of Technology (PL))

Presenter: GRACZYKOWSKI, Lukasz Kamil (Warsaw University of Technology (PL))

Session Classification: Statistical Methods for Physics Analysis in the XXI Century

Track Classification: H. Statistical Methods for Physics Analysis in the XXI Century